

111-A APPLICATION BASED ON:

Provisional Application Serial No. 60/192,400, filed March 27, 2000

Docket No:

80,525

Inventors:

Andrew Gallagher

Attorney:

Stephen H. Shaw

**A METHOD OF ESTIMATING AND CORRECTING CAMERA
ROTATION WITH VANISHING POINT LOCATION**

Commissioner for Patents
Attn: Box Patent Application
Washington, DC 20231

Express Mail Label No: *EM51164030US*

Date: *September 15, 2000*

005160-091500

**A METHOD OF ESTIMATING AND CORRECTING CAMERA
ROTATION WITH VANISHING POINT LOCATION**

CROSS REFERENCE TO RELATED APPLICATIONS

5 Reference is made to U. S. Provisional Serial No. 60/192,400, filed
March 27, 2000, entitled *Correction of Small Image Rotations Based on
Vanishing Point Detection* by Andrew C. Gallagher.

FIELD OF THE INVENTION

10 The invention relates generally to the field of image processing.
More specifically, the invention relates to estimating and correcting for
unintentional rotational camera angles that occur at the time of image capture,
based upon the captured image's corresponding location of its vanishing points.
Furthermore, the invention relates to image warping; the process of warping an
15 image in a manner that causes the vanishing points of the image to relocate.

BACKGROUND OF THE INVENTION

Conventional consumer cameras (both film and digital) capture an
image having a rectangular imaging area. For cameras using 35mm film, the
20 horizontal dimension of the imaging area is approximately 36mm and the vertical
dimension of the imaging area is approximately 24 mm.

Many photographers find it very difficult to hold a camera level
when capturing a photograph. A photographer often pivots the camera slightly
about the optical axis (whereby the optical axis is defined as an axis passing
25 through the center of the image capture frame and the center of the camera lens).
These small rotations are generally unintentional and may occur because
conventional consumer cameras are lightweight, for example, the popular Single-
Use cameras. Rotations about the optical axis give the appearance that the subjects
in the photograph are off-centered, leaning, or tilted.

30 U.S. Patent No. 6,011,585 issued January 4, 2000 to Anderson,
entitled *Apparatus and Method for Rotating the Display Orientation of a*

005160-091500

Captured Image describes a method of determining image format and orientation from a sensor present in the camera at the time of image capture. Whereas, this sensor can enable one to determine the orientation or format of a captured image, it cannot lead to detecting a small amount of camera rotation. Instead, the sensor
5 identifies the major image orientation (in increments of 90 degrees) by determining which side of an image corresponds with the "top" of the image. In addition, the Anderson method necessitates a sensor of this sort be present within a camera.

Lutton et al. (in "Contribution to the Determination of Vanishing Points Using Hough Transform," *IEEE Trans. Pattern Analysis and Machine Intelligence*, Vol. 16, No. 4, pp. 430-438, Apr. 1994) attempts to detect the
10 vertical direction of an image. The Lutton et al. article teaches one to select the direction that is orthogonal to the most directions in the scene. The implicit assumption is that the scene will contain many horizontal lines. However, this is not always the case. Additionally, the Lutton analysis is performed with a
15 possibly large number of line segments, rather than vanishing points. The Lutton method requires a great deal of processing and may be computationally costly.

Some existing systems detect and correct a skewed image. These systems are primarily in the field of document imaging. Nevertheless, documents have a fairly well defined structure, as opposed to more unpredictable consumer
20 type photographs. Another sharp contrast between document imaging correction and photographic imaging correction is that vanishing point detection is not used in the field of document imaging, nor is vanishing point detection required.

Pending U.S. Application Serial No. 09/281,574, filed March 30, 1999 by Gallagher, entitled ***A Method for Modifying the Perspective of a Digital Image***, shows a process for compensating from the perspective of image capture.
25 However, this solution cannot lead to compensating for camera rotation about the optical axis. The perspective of image capture relates to tipping the camera in the up and down direction (i.e. about the *x*-axis), which is much different than rotating a camera about the optical axis. Therefore, pending U.S. Application 09/281,574

005160-95029960

Consequently, a need exists for overcoming the above described drawbacks. More specifically, a need exists for determining slight image
5 orientation changes.

The need is met according to the present invention by providing a method of generating an image transform for modifying a digital image, that includes the steps of: detecting a vanishing point related to the selected image; 10 determining a preferable vanishing point location, and generating an image transform based on the vanishing point location and the preferable vanishing point location.

20 The present invention has an advantage of improving the method of
correcting for small angles of camera rotation, i.e. camera tilt.

The subject matter of the invention is described with reference to
25 the embodiments shown in the drawings.

Fig. 2 is a schematic of the reference systems employed to describe
30 the present invention;

Figs. 3A-C show several examples of images with different formats and orientations;

Figs. 4A-B show the effects on an image as a result of small angle camera rotation;

5 Figs. 5A-B show an actual image that has been corrected by the method of this invention; and

Fig. 6 is a diagram useful in describing the present invention.

DETAILED DESCRIPTION OF THE INVENTION

10 In the following description, the present invention will be described in the preferred embodiment as a method. Those skilled in the art will readily recognize that the equivalent of such a method may also be constructed as hardware or software within the scope of the invention.

Fig. 1 shows a block diagram of the present invention. The
15 purpose of the present invention is to estimate the amount of rotation of a camera relative to the scene at the time of image capture, based upon a digitized representation of an image. The source of the digitized image is irrelevant. The digitized image may be a scan of a film negative, a scan of a photograph, or an image captured with a digital camera. It should be well understood that in cases
20 where the digitized image is a scan of a hardcopy image that the rotation of the digitized image corresponds to the rotation of the source image. That is, if a photograph was captured with a camera that was tilted by β degrees, then the corresponding scan of that photograph (or the scan of the photographic negative) will also be tilted by β degrees. Thus, the method described herein may be used to
25 automatically determine the amount of camera rotation at the time of capturing the photograph, for example; by first digitizing the photograph and then analyzing the resulting scan with the method of the preferred embodiment. In this case, the source image is the photograph and the digital image is a result of the scan. It should be further understood that the source image may for example be a large
30 resolution digital image. This source image may then be decimated to generate

the digital image that the method of the present embodiment operates upon.

Again, the result from the present invention applies to both a source image and a digital image. Note that the preferred embodiment is described with reference to digital images having a dimension of 512 pixels by 768 pixels, although those skilled in the art will recognize that many image resolutions may be utilized with equal success. For example, the present invention is also suitable for processing a digital image having dimensions of 512 pixels by 512 pixels. Note also that the terms "digitized image", "digital image", and "image" are used synonymously throughout this description.

Vanishing points are a useful feature for determining the amount of rotation of an image because of the many vertical parallel lines associated with human construction. Despite a left to right positioning of the camera (or other image capture device), the vanishing point associated with the vertical scene lines nearly always falls near the vertical axis of the image.

Fig. 2 shows a reference system employed with the present invention. A focal point 2, representing the approximate location of the focal point of the lens used to capture the image, is located a distance f from the image plane 3. The focal point 2 represents the center of a Gaussian sphere. The x -axis and y -axis define the dimensionality of the image plane 3. The z -axis is also defined as the "optical axis" of the system. The x' -axis and the y' -axis define a plane that is parallel to the image plane 3. The image origin 4 is defined as the point of intersection of the image plane with the optical axis, and is given in Cartesian coordinates as $(0,0,f)$. Generally, the image origin is assumed to be at the center of the distribution of pixels comprising the digital image, although this assumption may not be correct. In fact, if the camera is constrained to only pivot up or down (pivoting solely about the x -axis), the vanishing point associated with vertical scene lines must fall on the vertical axis of the image. For many photographers, a pleasing captured image results by maintaining a level camera during shooting the image. Attempting to keep the camera level applies whether the camera is held by hand or placed on a tripod or some other mechanical device.

Because it is desirable to capture an image with a camera held level (i.e., pivoting solely about the x -axis), and such practice results in the vanishing point most closely associated with the image's vertical scene lines to fall on the vertical axis of the image, consequently, the vertical axis of the image is defined as a preferable
5 vanishing point location for achieving a user desired image composition. In other words, the captured image is preferably represented as a level depiction of the captured scene with respect to the vertical lines within the scene. However, if the camera is rotated about the optical axis (the z -axis), then the vanishing point associated with vertical scene lines will not fall on the vertical axis of the image.
10 However, the angle from the vanishing point corresponding to the vertical scene lines to the vertical axis of the image is equivalent to the amount that the camera was rotated about the optical axis.

In a human-made construction, there are also a large number of horizontal lines (lines orthogonal to the vertical lines). If these horizontal lines
15 happen to be parallel to the image plane (orthogonal to the optical axis), then a vanishing point for the horizontal lines will occur at infinity on the horizontal axis of the image. However, depending upon the position of the photographer, it is very likely that the horizontal scene lines will not be parallel to the image plane. If this condition exists and the camera is tilted about the x -axis, then the vanishing
20 point associated with the horizontal scene lines will not fall on either axis of the image plane. Thus, the vanishing point associated with horizontal lines in the scene may fall on the horizontal axis of the image, but because of the scene composition it is just as likely that the vanishing point will fall elsewhere. Hence, the vanishing point location corresponding to the horizontal lines in the scene is
25 not constrained to fall near an image axis, but it is highly likely that the vanishing point associated with the vertical lines of the scene will fall near the vertical axis of the image. Conversely, if a vanishing point is located near an image axis, it is far more likely that this vanishing point corresponds to a set of vertical scene lines than a set of horizontal scene lines.

Fig. 2 discloses an example of expressing a vanishing point location. First, a vanishing point may be expressed as a location on the image plane. The vanishing point v may be expressed as the point at location $v = (x_o, y_o, f)$. Such a representation of the vanishing point location performs well when the vanishing point is located near the image origin, however, the coordinate locations along the x -axis and the y -axis may grow quite large. Another aspect of the invention illustrated in Fig. 2 is a vanishing point representation scheme, commonly used in the field of vanishing point detection. In this representation, the vanishing point is represented with a Gaussian mapping. As such, each location v on the image plane has a unique location v_G on the Gaussian sphere. The relationship between the vanishing point location on the image plane $v = (x_o, y_o, f)$ and the vanishing point vector v_G determined by the Gaussian mapping is given with the equation:

$$v_G = \frac{v}{|v|}$$

Likewise, if the vanishing point vector $v_G = (x_G, y_G, z_G)$ is known and the vanishing point location in the image plane is desired, then the following projection is used:

$$v = \frac{fv_G}{z}$$

One can easily see that the vanishing point vector v_G is a unit vector extending in the direction from the focal point to the vanishing point location in the image plane. With reference to Fig. 2, the image plane is positioned f distance from the optical origin of the system. Normally f distance represents the focal length. If the focal length is unknown, then a reasonable guess may be used. In one embodiment, f is the diagonal length of the imager. For example, where an image has a dimension of 512 pixels by 768 pixels, f equals 923. The vanishing point on the image plane may then be represented as the unit vector that points from the optical system's origin to the vanishing point on the image plane. This vector is of length one and may be described as the coordinates of the intersection of a Gaussian sphere (a sphere of radius 1.0) centered at the

optical system origin (the focal point) and the line passing through both the optical system origin and the vanishing point on the image plane. This vector representation of the vanishing point is advantageous because it contains the entire space of the image plane.

5 In one embodiment of the invention as shown in Fig. 1, a digital image is input to a vanishing point detector 12. The purpose of the vanishing point detector is to identify the locations of the vanishing points of the digital image. As is well known in the art, a vanishing point is the result of the perspective projection of the three dimensional scene onto a two dimensional
10 image plane. A vanishing point refers to the point in the image plane (a two dimensional projection of the three dimensional scene) where parallel lines in the scene meet. Vanishing points generally only have relevance for images containing images of a structure containing at least two line segments, generally from man-made structures. Several authors have documented methods of automatically
15 locating the vanishing points of an image. For example, S. Barnard, "Interpreting Perspective Images," *Artificial Intelligence*, vol. 21, pp. 435-462, 1983. A preferred automated technique is disclosed in pending U.S. Provisional Patent Application Serial No. 60/192,195 filed March 27, 2000 by Gallagher, entitled ***Vanishing Point Detection by Training with Ground Truth Data***, which is
20 incorporated by reference. In addition, the vanishing point detector may include manually identifying the vanishing points using operator input. The vanishing point detector 12 outputs the locations of all vanishing points identified for the image. Characteristically, the number of vanishing points determined for a single image is not greater than three, although this should not be viewed as a limitation.
25 As described below, the vanishing points determined for the image are output in the form of vanishing point vectors. Assuming that M vanishing points are determined for the digital image, the output of the vanishing point detector may be represented as v_{Gm} , where m ranges from 1 to M . Alternatively, it is possible that zero vanishing points may be determined for the digital image. In this event,

then the present invention cannot draw any conclusions regarding image rotation from the location(s) of vanishing point(s).

If $M > 0$, then the vanishing point vectors detected by the vanishing point detector **12** are input to a vanishing point selector **13**. The purpose of vanishing point selector **13** is to determine those vanishing points that may be useful for determining the amount of rotation (i.e. the rotation of the camera from the level position at the time of capture) of the digital image by using the information contained in the M vanishing point vectors.

The vanishing point selector **13** determines if any of the M vanishing points associated with the digital image may be used to determine the amount of rotation of the digital image. Those vanishing points which may be used to determine the format of the image are referred to as "rotation candidate vanishing points."

In the preferred embodiment, a rotation candidate vanishing point is any vanishing point having an associated vanishing point vector $v_G = [x_G, y_G, z_G]$ meeting the either of the following two conditions:

1. $|x_G| > T_I$
2. $|y_G| > T_I$

Values of T_I between 0.3 and slightly less than 1 are useful in practicing the present invention. In the preferred embodiment, T_I is set at 0.5. The regions in the x_G, y_G plane in which such vanishing points lie are illustrated by cross hatching in Fig. 6. The vanishing point selector **13** outputs N (where N is between 0 and M) format candidate vanishing points.

Alternatively, it should be well understood that those skilled in the art can easily modify the above stated rules from classifying a vanishing point vector v_G as a rotation candidate vanishing point. For example, a vanishing point vector v_G may be classified as a rotation candidate vanishing point if the following condition is met:

$$\sqrt{x_G^2 + y_G^2} > T_2$$

Values of T_2 between 0.3 and slightly less than 1 are useful in practicing the invention.

If an image has zero rotation candidate vanishing points (if $N = 0$), then the present invention does not detect the rotation of the digital image using a
5 feature based on vanishing point location.

If there is at least one rotation candidate vanishing point corresponding to the digital image, then the rotation candidate vanishing point is output from the vanishing point selector 13 and input to a rotation detector 14. Note that other information may also be input to the rotation detector 14 in order
10 to aid the determination process. For instance the format (an identification of the vertical axis of the image as will be described herein) or orientation (an identification of the top of the image) may also be input to the rotation detector 14 for the purpose of determining the rotation amount. The operation of the rotation
15 detector 14 is to determine the angle β between the vertical axis of the image and the vector pointing from the image origin to the vanishing point corresponding to the vertical lines in the scene. In other words, the angle β corresponds to the amount of rotation of the capture device (for example a camera) about the optical axis.

The angle of rotation for each of the N rotation candidate vanishing
20 points is computed by the rotation detector 14 by first projecting the vanishing point under consideration to the xy -plane. This is accomplished according to the following equation:

$$v_{Gxy} = \frac{[x_G, y_G, 0]}{\sqrt{x_G^2 + y_G^2}}$$

The vector v_{Gxy} is a vector which generally points in a direction
25 parallel to the vertical axis of the scene.

Next, the angles γ are determined to be the angle between the vector v_{Gxy} and both the positive and negative vertical axes of the image. The vertical axis of an image is the axis on the image plane parallel to either the x -axis or the y -axis which also passes through the "top" and "bottom" of the image. The

Sub
2.1

vertical axis of the image will be further explained herein below. If the vertical axis of the image is known, then the angles γ are computed by taking the inverse cosine of the dot product of the two vectors, as is well known in the art. For example, if the y -axis is the vertical axis of the image, then the angles γ may be determined as:

$$\gamma = \text{sign}(x_G y_G) \cos^{-1}(v_{Gxy} \cdot [0, 1, 0])$$

$$\gamma = \text{sign}(x_G y_G) \cos^{-1}(v_{Gxy} \cdot [0, -1, 0])$$

where $\text{sign}(x_G y_G)$ represents the sign (-1 or +1) of the product of x_G and y_G . For example, when x_G and y_G are either both negative or both positive, the $\text{sign}(x_G y_G) = 1$. Alternatively, when only x_G or y_G is negative then $\text{sign}(x_G y_G) = -1$.

If the vertical axis of the image is not known by the rotation detector 14, then the angles γ are determined to each image axis (for a total of four angles γ). Thus, for each rotation candidate vanishing point, several angles γ are determined as the angle between the vector v_{Gxy} representing the projection of the vanishing point onto the xy -plane and the image axes. The angle β is determined to be the smallest (in an absolute value sense) of these angles γ . Note that the angle β may be positive or negative in magnitude, indicating the direction of camera rotation. The output of the rotation detector 14 is the angle β . Note that the angle β is expected to be equal to or less than $\pi/4$.

Fig. 1 shows that the identification of the vertical axis of the digital image is output from the rotation detector 14 and passed to an image transform 19. According to an alternative embodiment of the invention, this image transform may also receive various other parameters, including the digital image itself. The operation of the image transform 19 may be any number of image transformations that benefit from knowledge of image rotation.

In an example of the use of the rotation information, the angle β of the digital image may be stored as metadata by the image transform 19. Metadata is generally defined as data corresponding to an image apart from actual pixel values.

Additionally, the function of the image transform **19** may account for the rotation of the image by performing an image rotation in the opposite direction. Image rotation is well known by those skilled in the art of image processing. The amount the digital image must be rotated is given by the negative of β .

In another embodiment, the vanishing point vector is interchangeable with the vanishing point location given in Cartesian coordinates.

Fig. 3 shows several image examples useful for defining terms used herein. Fig. 3A shows an example image of a person. The top of the image is defined as the side of the image that was in the “up” direction from the photographer’s perspective at the time of image capture. Identification of the top of an image solves the problem of orientation. In the case of Fig. 3A, the top of the image clearly corresponds to the top of the subject’s head. Note that the bottom of an image is always the image side opposite the image top. Also shown in Fig. 3A are axes parallel to the x -axis and the y -axis passing through the image origin. These axes have herein been defined as the x' -axis, and the y' -axis, respectively, and shall herein be known collectively as image axes.

Note that the vertical axis of the image is that line that passes through the top and bottom of the image, as well as the image origin and coincident with either the x' -axis or the y' -axis. The vertical axis of the image defines the format of the image, by specifying which two sides are top and bottom. As used herein, the term “format” means the identity of the vertical axis of the image. However, the vertical axis of the image does not specify which of the two sides is the top, thus the orientation of an image may remain unknown even when the format is known. Note in the case of the image in Figure 3A, the y' -axis is the vertical axis of the image.

Fig. 3B shows an example where the top of the image is the right image side. In this example, the vertical axis of the image is the x' -axis. Finally, Fig. 3C shows an example image where, like the image in Fig. 3A, the vertical

axis of the image is the y' -axis. However, in this example the top of the image lies on the negative y' -axis.

Note that if the camera is not held level at the time of image capture, then the image may appear to be tilted. That is, if the x -axis of the imaging system is not orthogonal to the vertical direction in the scene, then the image will appear to be tilted or rotated. Commonly, the amount of rotation is small since the photographer is making an effort to hold the camera level. For example, Fig. 4A shows an embodiment similar to that of Fig. 3A, except for the fact that a simulated amount of camera rotation is included. Additionally, Fig. 4B discloses an embodiment to that of Fig. 3B with the addition of camera rotation. It is possible to show that the amount of the camera rotation (the angle from the level position) is the same as the angle β that the image is rotated. As described herein, the vanishing point locations detected from the image allow for the identification of this angle.

Fig. 5A shows an image that has an apparent amount of camera rotation (the camera was tipped in the clockwise direction at the time of capture.) The method of the present invention was applied to the image and estimated that the amount of rotation of the image in the counter-clockwise direction was 3.4 degrees. Fig. 5B shows the corrected image, generated by rotating Fig. 5A by 3.4 degrees in the clockwise direction. There is a noticeable improvement in the handrail orientation as a result of the processing of the present invention. In the case where the image transform 19 performs a rotation of the digital image, there may be additional logic based upon the value of β . For instance, if $|\beta| < 1$ degree, performing a rotation correction may not produce a result noticeable enough to be worth the computational effort. Additionally, there may be an upper limit on the amount of rotation that the image transform 19 will execute. For example, if $|\beta| = 44$ degrees, it may be advantageous that nothing is done to the image, as a modification of this magnitude may produce displeasing results if the value of β produced by the algorithm was incorrect. Another aspect of the invention has the image transform 19 performing a rotation by an amount of $-\beta$, the resulting image

005160-950E950

Sub
20

produced has a vanishing point on the vertical axis of the image (assuming that the value of β is correct). In this embodiment, the image transform 19 is a rotating transformation which operates by rotating the image. Such a transformation is a warping of the image, since the geometry of the image output from the image transform 19 has been modified relative to the geometry of the image input to the image transform 19. The location of the rotation candidate vanishing point v_G of the image input to the image transform 19 is an undesirable vanishing point location, because it does not lie on the vertical axis of the image. The vertical axis of the image is considered to be a preferable vanishing point location. However, those skilled in the art will recognize that other preferable vanishing point locations may exist. For instance, infinity (or $z_G = 0$ in vanishing point vector notation) may be a preferable vanishing point location. In this case, the digital image may be modified by an image transform 19 designed to warp the image in such a manner that the image resulting from the image transform 19 has a vanishing point now located at infinity. Alternatively, another preferable vanishing point location may be at infinity and also on the vertical axis of the image (in this case, either x_G or $y_G = 1$). Clearly, the vanishing point of the image output from the image transform 19 lies on the vertical axis of the image, and is therefore a preferable vanishing point location. Thus, the operation of the image transform 19 is to warp the image in such a fashion that a vanishing point associated with the input image migrates from an undesirable vanishing point location to a desirable vanishing point location in the image output from the image transform 19. Those skilled in the art of image processing will recognize that an image transform 19 may be created (either deterministically or empirically) to warp an image in such a manner as to relocate a vanishing point from an undesirable vanishing point location within the input image to a desirable vanishing point location within the resulting image.

Those skilled in the art will recognize that many variations may be made to the description of the present invention without significantly deviating from the scope of the present invention.

2	focal point
3	image plane
4	image origin
12	vanishing point detector
13	vanishing point selector
14	rotation detector
19	image transform